DOCUMENT RESUME

ED 104 815 SP 009 048

AUTHOR Soar, Robert S.
TITLE Validation of Competencies.
PUB DATE 31 Mar 75
NOTE 7p.; Paper presented at the Annual Meeting of the American Educational Research Association (Washington, D.C., March 1975)

EDRS PRICE MF-$0.76 HC-$1.58 PLUS POSTAGE
DESCRIPTORS Academic Achievement; Educational Objectives; Evaluation; Performance Based Teacher Education; *Performance Criteria; Personal Growth; Research Projects; *Student Behavior; *Teacher Behavior; *Teaching Skills

ABSTRACT

Teacher behavior is not likely to be "valid" in general, but rather only under specified conditions, such as the nature of the learning objective which is being pursued, the personal and demographic characteristics of the pupil being taught, the length of time during which learning takes place, and probably the context of other classroom behaviors in which the given behavior is experienced. The need for specifying what, for whom, how long, and what else, in terms of the application of validity specifications, are recognized in the Georgia project. The type of achievement measures selected and data obtained to meet this need call for an analysis which recognizes the influence of multiple measures simultaneously. In the Georgia project, data analysis will therefore be by multiple regression and, where adequate degrees of freedom are available, by multivariate analysis. The project will attempt to minimize the problem of inconsistency by replicating results over several years, sharpening hypotheses, and refining measurement procedures. Although past work in measuring teacher behavior and relating it to measures of pupil growth has not produced extensive agreement on findings, the improvements in methodology which are needed have become clearer. More rapid progress may now be possible where past progress has been slow and uncertain, and the Georgia project will be in a position to profit from these advances. (PB)

ED104815

U S DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

VALIDATION OF COMPETENCIES*

Robert S. Soar
College of Education
University of Florida
Gainesville, Florida

Although the major focus of this paper is validation of competencies, a word about the origin of the competencies studied in the Georgia project seems in order. Competency lists which had been developed in a number of states were reviewed by a committee of Carroll County teachers and competencies were selected which they felt should be included for certification, and they were rank-ordered in relative importance. When reviewed by a larger group of teachers in Carroll County, agreement with the competency list and the rank-order was high.

After the competencies had been identified, existing observation instruments were reviewed as sources of behavioral measures. Five observation instruments with relatively extensive research histories were selected for use, from which behavioral measures of numbers of the competencies could be abstracted, although it was not possible to represent all competencies. Data collection with those instruments is currently underway in classrooms.

A central element in the rationale for both competency selection and validation has been the assumption that a competency-based certification system could ultimately be expected to stand court-test, so that the data on which certification was based needed to be as objective as possible. This meant that "counting measures" should be strongly preferred, if not obligatory, and that establishment of the relations between these measures of teacher behavior and measures of change in pupils--validation--was centrally important. Recent research in teacher effectiveness indicates that there are a number of challenges which this validation effort must meet if it is likely to be meaningful.

SP009 045

---

*Presented in a symposium titled "Mandated Competency-Based Teacher Certification and the Public Interest", American Educational Research Association, Washington, D.C., March 31, 1975.

Validation of Competencies

Validity Specification

An over-all orientation which is developed in more detail elsewhere (Medley, Soar and Soar, 1975) is that a teacher behavior is not likely to be "valid" in general, but rather to be valid only under specified conditions--such as the nature of the learning objective which is being pursued, the personal and demographic characteristics of the pupil being taught, the length of time during which the learning takes place, and probably the context of other classroom behaviors in which the given behavior is experienced.

The Nature of the Objective - It does not seem probable that the same teacher behavior is best for pupil growth in all objectives. That is, the teacher behavior which facilitates complex cognitive growth such as abstracting, generalizing, inferring, or creative endeavors is probably different from that which facilitates memorizing facts such as the multiplication table or the day's list of fifteen spelling words. Several sets of data support this conclusion in varying degrees (Solomon, Bezdek and Rosenberg, 1963; Soar, 1968; Soar and Soar, 1972; Soar and Soar, 1973). If this is true, or unless we are sure it is not true, it seems important to collect achievement measures which represent varying degrees of cognitive complexity and measures of non-cognitive change as well, against the possibility that desirable changes for one may not parallel those for the others. It would follow, then, that observational data should be recorded separately for the teaching of different objectives.

Pupil Differences in Response to Classroom Behavior - It does not seem probable that the same teacher behavior is best for all pupil groups. Recent evidence suggests that a teacher behavior which is best for high socioeconomic status pupils is not always best for low socioeconomic status pupils. At least three recent sets of data agree in indicating that the emotional climate of the class-

room which is associated with greatest achievement growth is different for pupils who differ in socioeconomic status (Brophy and Evertson, 1974; Soar and Soar, 1973). Scattered findings suggest that other characteristics of pupils may also be related to the kind of classroom experience which is most functional for them.

The Time Period for Pupil Growth - Although relatively short time periods, ranging from a few minutes to a few weeks, have been advocated as a basis for validating competencies, generalizing from short to longer time periods seems risky without empirical evidence. It seems probable that it will be easier to show the growth of relatively low level cognitive skills in a short time while longer time periods would be needed to show growth in more complex cognitive skills. There are suggestions that the teacher behavior which is associated with pupil achievement gain during the school year may be different from the teacher behavior which is associated with pupil gain the following summer while out of school (Soar and Soar, 1973) and there is convincing evidence that practically important amounts of pupil growth (as measured by change in standardized achievement test scores) occurs over the summer while pupils are out of school (Hayes and Grether, 1969; Soar and Soar, 1969).

The Context of Other Behaviors - It seems intuitively reasonable that the impact of any single teacher behavior on pupils will be moderated by the context in which it occurs. Some evidence indicates that this is the case (Soar and Soar, 1973). For example, the frequency with which the teacher took a central role in the activities of the classroom was unrelated to any pupil outcome measure by itself, nor was the frequency with which the pupil's learning task had been specified by the teacher. But when the two behaviors were examined simultaneously, frequent occurrence of one or the other was associated with increased pupil achievement gain, but frequent occurrence of both, or infrequent occurrence of either, was associated with decreased pupil gain. Numbers of other such interactive relationships have been found.

Another example of impact being moderated by context occurred when a measure of classroom behavior with a near zero correlation with pupil change became significantly related to pupil change when other measures of classroom behavior were held constant statistically. This result presumably parallels the expectation in experimental research that the significance of the experimental variable can only be tested meaningfully when all other variables are held constant.

Application of Validity Specifications - The need for specifying what, for whom, how long, and what else, indicated by the results of past research, are recognized in the Georgia project. The staff has selected achievement measures which differ in complexity, as well as a limited set of non-cognitive measures; they are collecting data on the socioeconomic status of individual pupils; they are collecting pupil data for the school year as well as for the summer in order to relate both to the pupil's classroom experience, and they will examine interactive relationships and the possibility that some teacher behaviors relate to pupil outcomes only when others have been held constant.

The Need for Complex Analysis

This plan clearly implies the need for an analysis which recognizes the influence of multiple measures simultaneously. If interactions between classroom behaviors are to be tested, and if interactions of classroom behavior with pupil personal and demographic characteristics are to be tested, and if the effects of some classroom behaviors are to be examined while holding other classroom behaviors constant, the need for a complex analysis is clear.

Another aspect of the data will need to be dealt with by the analysis. This is the probability that numbers of classroom behaviors, and possibly some pupil characteristics as well, will be related in nonlinear fashion with pupil outcome measures (Solomon, Bezdek and Rosenberg, 1963; Soar, 1968; Coats, 1966; Soar and Soar, 1972; Soar and Soar, 1973; Brophy and Evertson, 1974). This also seems

eminently reasonable, once the question is raised. If some of a teacher behavior is good, will increasing amounts of it be better, without limit? And yet when a product-moment correlation is calculated, this assumption is implicitly made. So it seems clear that the data analysis will need to recognize the simultaneous effect of considerable numbers of variables, the interactions of numbers of variables, and also the possibility of nonlinear relationships between variables. To complicate things further, the ultimate--nonlinear interactions--have also emerged in past work (Anderson, 1970).

In the Georgia project, analysis of the data by multiple regression is planned, which will permit testing multiple variables simultaneously, along with interactions, nonlinear relations, or nonlinear interactions (Kerlinger and Pedhazur, 1973). Where adequate degrees of freedom are available, multivariate analyses may also be employed (Darlington, Weinberg and Walberg, 1973).

The Problem of Power

The limited power of the typical study of teacher effectiveness has not generally been recognized until recently. Cohen (1969) argues convincingly that in a study which assumes a relation of .30 between two true measures, a reasonable degree of fallibility in measurement procedures would reduce the observed relation to .10. In a study with an N of 15 classrooms, which has not been uncommon, only 6 out of 100 relations which are real would be found to be significant. Much of the inconsistency of past results presumably follows from this lack of power.

The Georgia project will attempt to minimize this problem by replicating results over several years, sharpening hypotheses and refining measurement procedures.

A Concluding Comment

Although past work in measuring teacher behavior and relating it to measures of pupil growth has not produced extensive agreement in findings, the improvements in methodology which are needed have become clearer. Progress may now be possible

where past progress has been slow and uncertain, and the Georgia project will be in a position to profit from these advances.

The critical issue for certification in particular, and education in general, is that empirical validation is a central aspect of the ongoing development, in contrast to programs, however innovative, behavioral, or whatever, in which there is no concern for evidence that the effort makes any difference to pupils. That is an issue whose importance can scarcely be overstated.